EXHIBIT 3

OAU4AUT1

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK
----------x

AUTHORS GUILD, *et al.*,

Plaintiffs,

v. 23 Civ. 8292 (SHS)OTW

OPENAI, INC., *et al.*,

Conference

Defendants.

----------x

New York, N.Y.
October 30, 2024
9:40 a.m.

Before:

HON. ONA T. WANG,

U.S. Magistrate Judge

APPEARANCES

SUSMAN GODFREY LLP
Interim Class Counsel for Authors Guild and Alter Class Actions
BY: ROHIT NATH
CHARLOTTE LEPIC
JORDAN CONNORS

SUSMAN GODFREY LLP
Attorneys for The New York Times
BY: EMILY CRONIN
DAVIDA BROOK
ELLIE DUPLER
ZACH SAVAGE
ALEXANDER P. FRAWLEY

ROTHWELL FIGG
Attorneys for New York Times and Daily News
BY: STEVEN LIEBERMAN
JENNIFER MAISEL
KRISTEN LOGAN

Appearances continued

LOEVY & LOEVY
Attorneys for Center for Investigative Reporting
BY: MATT TOPIC
RANDALL W. JACKSON

LIEFF CABRASER HEIMANN & BERNSTEIN, LLP
Attorneys for Authors Guild and Alter Class Plaintiffs
BY: RACHEL GEMAN
WESLEY DOZIER
ANNA J. FREYMANN

KEKER VAN NEST & PETERS
Attorneys for OpenAI
BY: NICHOLAS GOLDBERG
PAVEN MALHOTRA
MICHELLE YBARRA
ANDREW DAWSON

MORRISON & FOERSTER LLP
Attorneys for OpenAI
BY: JOSEPH GRATZ
CAROLYN HOMER

COWAN DeBAETS ABRAHAMS & SHEPPARD LLP
Attorneys for Authors Guild and Alter Plaintiffs
BY: SCOTT J. SHOLDER

LATHAM & WATKINS LLP
Attorneys for Microsoft
BY: ANDREW GASS
ELANA NIGHTINGALE DAWSON

FAEGRE DRINKER BIDDLE & REATH LLP
Attorneys for Microsoft
BY: JEFFREY JACOBSON

(Case called; appearances noted)

THE COURT: All right. Let's see. I think I tried to create my own agenda to try to get through matters a little bit more quickly than we did last time. I believe I may only have blocked two hours for you and hope not to go over.

All right. I want to start with something that shouldn't be controversial, so it's not an invitation for you to make an argument, but there is a motion to consolidate the Center for Investigative Reporting case with the newspaper cases. Can someone confirm -- there is no disagreement about whether consolidation should happen; right?

MR. TOPIC: I was told to go to the microphone. Good morning, your Honor. Matt Topic on behalf of Center for Investigative Reporting.

No objection in concept. But definitely an objection to the 80-day fact discovery period that would result under the schedule that's existing.

THE COURT: We are not doing that.

Next?

MS. BROOK: Good morning, your Honor Davida Brook on behalf of the Times. No objection. But we hope that the schedule does not get kicked out too long because that would be prejudicial to the Times.

THE COURT: Other than objections about the schedule, are there any objections to consolidation?

MR. TOPIC: Not from us.

MS. MAISEL: No objection from the Daily News.

MR. MALHOTRA: Your Honor, Paven Malhotra from OpenAI. We obviously support consolidation.

THE COURT: As you can tell, I'm loud enough without the microphone. If you are going to come up to the podium, make sure we can all hear you. You know what, I'm not going to undo any direction from my deputy because she is right, and I'm not right. I can hear you better if you speak at the podium and plan to use the microphone. I guarantee you will not be louder than me.

So there is no objection to consolidation.

We will deal with the schedule right now. Let me take a look at the scheduling. I'm going to grant consolidation since there is really no material objection.

All right. And then as for the timing, I know the Center for Investigative Reporting expressed concern about the fact discovery deadline. The December 20, 2024, deadline was always intended, and it was explicitly called in interim discovery deadline. We know we are going to push it out. But I put that in there as a little bit of a benchmark for myself as well to just make sure that discovery is moving. That doesn't mean that we're moving your ultimate deadline up at all. Okay. We are going to try to get through these cases as well as we can.

One thing I will import from the Center for Investigative Reporting's schedule is the last day to amend pleadings or join additional parties without leave of Court is January 8, 2025, which I believe is the date that you had anyway.

As far as interim discovery deadlines that are agreed upon between the parties, I tend to try to avoid micromanaging that. You are free to agree. You are free to push those out. What I want to see is an understanding of where either the actual fact discovery deadline or the interim fact discovery deadline is and whether that needs to be pushed out a little bit. In a case like this, we will probably have to take it month to month. So by November or certainly by December, we will have a better idea of how far out to push that December 20th interim deadline. I won't say more than that. Because given the amount and the nature of the disputes that seem to be percolating upward, I'm not sure if by December that we will have any better idea of what the ultimate discovery deadline is.

Ms. Brook, you are standing up. You have something you want to say about that?

MS. BROOK: If it pleases the Court, just one suggestion. The Times was thinking that this was a likelihood today, so we have the following suggestion, as your Honor said, to keep us all moving with the recognition that dates may need

to move some which is to keep the current December 20th deadline as the cutoff for trying to get all our documents to each other, with the recognition that depositions will likely extend beyond that into 2025.

THE COURT: I have no problem with that. Like I said, I just said, interim deadlines I take less of a problem with. I'm not going to take a position on other dates that are proposed in 2025 or substantial production dates. Right now those are too far out. I have also had other cases where the parties end up fighting about what constitutes a substantial production by that date, and you all have enough disputes without me needing to add more. But I agree with you, Ms. Brook, that we will keep December 20th and see where we are in terms of document production.

I will also give you a little bit of a broader, general guidelines which is going to bleed into comments about the general coordination of discovery. My hope is that I will be able to settle some of the multiple filings that happened over the last two or three weeks. And I don't mean settle in cases like a settlement conference, but we need to turn it down a little bit. Because we are getting inundated with so many words and so many filings that it's hard to keep sight of what we need to be dealing with right now. We are still fairly early in the case.

I'm not going to require anybody to formally

coordinate with the actions in the Northern District of California. What that means is that I'll keep a close watch on timing, but I don't expect that this case should be used to slow anything down in the Northern District of California, that case is going to go on its own timetable. I also will look long and hard if something happening in the Northern District of California is being used to try to slow things down here. All right. There are cases here, and they are going to move. As I understand it, and having read the opinion, we are on different timelines. We're on different tracks. There are pending motions to dismiss here that haven't been decided. So we are in a different place, and we are moving on discovery.

And that brings me to my general principles that I want to talk about today. I am all in favor of cross production of documents, of sharing or cross production of testimony for efficiency purposes. What that involves -- and this is not a limiting example, this is just an example. You don't have to ask the same pedigree questions over and over again. You don't need to ask basic background questions over and over again and document requests, interrogatories, or depositions. You don't have to make a general requests and redo searches and production. However, prior production should not be used as an "ah ha" opportunity to prevent, preclude, or delay inquiry into otherwise relevant professional areas. I will impose costs under 37(a)(5) if I find that a party is

using cross production or coordination in this way.

I really encourage you all to be flexible and be creative here. For example, I recall at the last conference I had suggested perhaps that some early 30(b)(6) depositions could help focus the document discovery in other depositions. I agree and I understand that's not normally how 30(b)(6) depositions are used. Usually they are used at the end of discovery but in a case like this where there are so many sub issues so hard fought and where, to some extent, some of the documents discovery issues relating to the training of the defendant's generative AI, it might be productive to have a broader overview 30(b)(6) that's without prejudice to future depositions. And similarly, if there's a cross 30(b)(6) for plaintiffs that might get to some of the early issues that seem to be holding up some of the discovery. That said, again, I don't expect my words to be used to be like, "the Judge said we could have a 30(b)(6) deposition." I really want to encourage you to think about whether that can cut through some of the documents disputes we are having and some of the custodian disputes.

All of your proposals and percolating disputes about limitations, depositions or coordinated deposition protocol, I think they are premature right now. We are still fighting about documents. I also don't think that now, while we are in the midst of document production, that we are still disagreeing

about what's relevant and proportional to produce, there shouldn't be a time to limit the numbers or length of depositions. You're disagreeing about custodians for document production. I'm not sure how that enables anybody to make an informed decision about who should or shouldn't be deposed or for how long or how many times.

What I will tell you is that I will follow Rule 30. I will not presumptively limit witness depositions. Again, I encourage you to work together and to tailor what you learn to tailor your depositions and your deposition plan from what you've learned from document review just like any other case. It's just a big case. I don't want to invite multiple motions to reopen or extend deposition time. But at the same time, if the parties can't agree, I will take a look at that. Because I understand that we are still at a point where the parties may learn things during a deposition or at a deposition that may result in the need to recall a witness for more deposition or extend some time for another witness because it suddenly turns out that somebody finds that there was somebody who are more involved than they thought there were, for example.

General guidelines when I have done this before is if there is a motion, it's got to follow the principles under Rule 30 which also means I need to see the whole deposition, not just the snippets of the deposition so I understand the context.

The parties that are seeking more time or more information has to explain what it is that they are looking for and why and why it came up, and why you aren't able to finish in the time that you had.

And I really hope that the parties can agree that if there is a need that you could do this and you can agree to do this without motion practice. Because if there is motion practice, I will start looking at apportioning or shifting costs under Rule 37(a)(5).

All right. So this brings us to document production issues. Based on what I just said with the depositions, I see that there is a dispute about custodians. Is that a ripe dispute? Do you think it's a ripe dispute? Okay. When you are going to argue it, point me to the ECF number in the case, and give me a moment to catch up. One thing I understand is that it's defendants who are objecting to more custodians. Is there a custodian issue on the other side also?

MS. YBARRA: Your Honor, Michelle Ybarra from OpenAI. There is, but it's not ripe for the Court today.

THE COURT: So let's address this custodian issue.

MR. NATH: Your Honor, one point of clarification. There are two custodian issues teed up for the Court today, one in the class plaintiff case and then one in the New York Times case. I think New York Times is going to handle theirs first.

THE COURT: So which ECF Number in which case is this?

MR. FRAWLEY: Alexander Frawley for the New York Times. This is Document 204 in the New York Times case.

THE COURT: Okay. Let me find it. It doesn't help when my summary was 29 pages long. Okay. Bear with me. Go ahead.

MR. FRAWLEY: Thank you, your Honor. Just as an update, last time we were here, and I was in front of you on this motion, your Honor's decision was that OpenAI should preserve the documents for the eight custodians in dispute. And that then we would review their forthcoming custodial productions and revisit whether we still want and need these eight custodians. As an update, since then -- that was September 12 -- we didn't receive any new custodial productions. There was a production last night from OpenAI. Maybe it has custodial documents. We don't know yet. But we were planning to review the forthcoming productions and be ready today to talk about any updates to the documents. We don't have any updates to our motion because we haven't received any custodial documents.

So we think our motion is ripe. In the motion we cite documents that we have received earlier on in discovery from OpenAI --

THE COURT: Okay. Back up just a minute, sorry.

MR. FRAWLEY: Sure.

THE COURT: So I said preserve the custodians, review

the productions meetings and meet and confer. You said you did not receive a custodial production. What does that mean versus what you gotten so far?

MR. FRAWLEY: So we did receive custodial documents earlier on in discovery.

THE COURT: Custodial documents from who?

MR. FRAWLEY: From both Microsoft and OpenAI.

THE COURT: I know. But when you say "custodial documents," you mean the agreed-upon custodians?

MR. FRAWLEY: I'll back up. I'm sorry, your Honor.

THE COURT: Okay.

MR. FRAWLEY: Initially OpenAI selected 12 custodians. They produced documents from those custodians' files earlier on in the summer.

THE COURT: And those are what you are calling "custodial documents"?

MR. FRAWLEY: Yes. I'm sorry.

We reviewed those productions, and we used those productions to decide who else we wanted to pick as custodians. And we cited many of those documents in the motion at Docket 204.

And then I think your Honor's instruction was we would be getting more custodial documents, so we should review those additional documents and revisit the eight that we had picked. My update is just that, until perhaps last night, we didn't

receive any additional documents that we could use to update our motion.

So we think our motion is ripe. We asked OpenAI recently if they had any revisions from their position, and they will correct me if I'm wrong, but I think their position is that they still disagree on all eight of these custodians.

So we ask your Honor to consider our motion. I'm happy to answer any questions you have about any of the eight or walk through them. I know time is tight, so we are happy to rest on the papers as well.

THE COURT: Point me to the ECF number of OpenAI's objection or opposition? Is that the one that actually lists out each one and then says why it's not proportional?

MS. YBARRA: Yes, your Honor. Michelle Ybarra for OpenAI. That's Docket Number 211.

THE COURT: So what's the objection? Is it that these additional custodians are not relevant or that production from these custodians would not be proportional?

MS. YBARRA: Yes, your Honor, both of those are correct. Counsel also, I believe, is misinformed about the status of our production. We did produce about 15,000 custodian documents this week. We have been working very hard on multiple fronts. We made significant progress in technical discovery, in producing training data, for collecting source code. We're making modules available for inspection --

THE COURT: But that's not what these custodians are about, right? These were higher-level people?

MS. YBARRA: That's correct, your Honor. These are -- I believe New York Times is seeking, you know, their email and custodial files, files in their possession. However, we designated 24 OpenAI former and current employees as custodians, including our most senior executives, and we have produced documents from all 24 of them.

So I know the New York Times still has not had an opportunity to review our most recent production. It was substantial, and we have substantial productions coming. We have a lot in the cue. I think the Court's guidance at the last hearing still applies. We should give plaintiffs an opportunity to review the production, even in what we've seen in our arguments so far, they are making demands for custodians who are duplicative of the existing individuals.

And to your Honor's point earlier about we need to move things forward --

THE COURT: Okay. So talk to me. I mean, these custodians, their documents have been preserved. What's the volume of their documents in total, either in data, size, or a number of documents and number of pages?

MS. YBARRA: For the custodians who are still in dispute, I don't have those figures.

THE COURT: If you are saying that they're duplicative

of what's already been produced, then presumptively they're relevant, so then we move on to proportionality. If the documents have already been segregated and preserved, and you are not prepared to make a proportionality argument, what --

MS. YBARRA: I'm absolutely prepared to make a proportionality argument.

THE COURT: Go ahead then.

MS. YBARRA: Let me back up and explain how we got here. Over the past several months, the New York Times has, at various points, demanded nearly 40 OpenAI custodians. They demanded some; they've dropped some and picked others. This whole process has been piecemeal and haphazard and contrary to the law. The New York Times' motion encapsulates that. That motion that they filed before the last hearing seeks production from 17 custodians. We've agreed to six of them. The New York Times dropped three of them. The remaining dispute is about these eight.

THE COURT: So let's back up a little bit. I mean, in a typical case, not one that has my courtroom full of people who really want to hear about document discovery disputes, a corporate defendant would just sort of say the usual sort of objection to a custom might be they don't have anything. So why are we bothering this person because they really don't have anything or are really low level. You know what, if it's going to be duplicative, knock yourselves out. Go for it. Why is

that not the case here?

MS. YBARRA: Because, your Honor, we are already inundated with hundreds of discovery requests, demands for more. We've had I think on average a meet and confer every business day since we were last before your Honor. We are reviewing documents at a breakneck pace and producing them.

I want to be clear, earlier your Honor asked what's the volume of the data of these custodians that's been segregated? I don't know the overall volume of the information that been preserved. But I can tell you based on the search terms that the New York Times requested for these new custodians that would add in the range half a million to another million documents for OpenAI to review. We are already underwater with the current existing custodians.

THE COURT: How is OpenAI reviewing these documents? Do you have an attorney laying eyes on everything, or are you using technology-assisted review?

MS. YBARRA: We are using contract reviewers and attorneys at the defense firms.

THE COURT: Is an attorney laying eyes on every document being reviewed?

MS. YBARRA: At some point, I believe an attorney is laying eyes on every document, yes.

THE COURT: I'm not going to be --

MR. FRAWLEY: Your Honor, may I be heard?

THE COURT: Go ahead.

MR. FRAWLEY: Thank you, your Honor. Alexander Frawley for the New York Times.

I know counsel for OpenAI just said that they think we should follow the process of reviewing documents and coming back, and I would just like to point out that that was the procedure that we had agreed to last time, but then we didn't get any documents in time to follow through on that procedure.

THE COURT: Yes, and I'm not happy about that, by the way. The night before is not sufficient time to prepare for something like this. And it's not sufficient time to say, you know what, this is premature because you've had over a month to do this.

Go ahead, Mr. Frawley.

MR. FRAWLEY: For these reasons, your Honor, we would ask you to rule on our motion at Docket 204. We think that would help move the case forward. And whatever the result is on custodian, at least we will have finality, and we can move forward on the productions.

THE COURT: I'll tell you what, I'm going to do something a little bit different. I want to make sure the defendants have the opportunity to be heard. So I'm going to direct defendants to use technology-assisted review to tell me of the volume of the eight custodians, the documents that have been segregated, what is the approximate amount of duplication,

of duplicative documents? And I'm sure you can come up with different ways to search fields to let me know either the amount of duplicative documents or the amount of documents for which no other current custodian is a custodian. Okay? And you can let me know, preferably both in terms of the amount of data, a measure of data, and also documents and/or pages.

Okay. How much time do you need to do that?

MS. YBARRA: Two weeks, your Honor.

THE COURT: Two weeks. All right. Take a look at the calendar.

THE DEPUTY CLERK: November 13.

THE COURT: I will give you till the 13th.

MS. YBARRA: Thank you, your Honor.

THE COURT: And you are going to file that on the docket. If there is something that needs to be redacted, you will follow the same procedure, make sure the plaintiffs get to see what it that is. Okay.

Mr. Frawley, anything more on that?

MR. FRAWLEY: Just two quick points, your Honor. The first is I would ask is, as part of this update, we get hit counts for the current search term proposal that we made. I know search terms are also on the agenda. I don't want to skip ahead, but I think that would be helpful.

THE COURT: Let me write down what I ordered and then let me hear a little bit about what you just said.

Okay. So for November 13 the first task, the first homework is for the eight custodians at issue, information regarding either duplicate document count that are duplicative of current custodians, not duplicative of the eight. It's exactly the two sets that you are comparing or unique documents.

And Mr. Frawley, you were asking for a hit count?

MR. FRAWLEY: Yes. We made a recent search term proposal. And as part of that back and forth meet and confer process, we are getting hit counts for the 26 custodians that are agreed to now. So I think it would make sense to include those hit counts for the additional eight. I didn't expect that to be controversial. I just wanted to make sure.

THE COURT: Why don't you do that too. Because that does help with the defendant's relevance arguments, so I appreciate you bringing it up now. If the hit count for the eight is really, really low, that might also play into the proportionality determination, right?

Yes?

MR. FRAWLEY: Sorry. Just one more quick point. Just to correct the record, I don't think we need to discuss anything further, but I know counsel for OpenAI referred to us or sort of characterized custodial production as "email." I imagine she was referring to the email in the shorthand way. We are seeking more than just email, but I don't think we need

to get into that right now.

THE COURT: I think much of this document is all electronic documents in ESI; right?

MR. FRAWLEY: Yes.

THE COURT: I know, sometimes I call things email.

MR. FRAWLEY: That's how I think about it.

MS. YBARRA: It was a shorthand with reference, yes, your Honor.

THE COURT: That's okay. All right.

What's the next ripe issues as far as documents?

MR. NATH: Your Honor, Rohit Nath for the class plaintiffs.

I just wanted to get some clarification in terms of what order to go in. There are some issues teed up from the class plaintiff side. There are issued teed up from the New York Times' side. The New York Times wanted to deal with theirs first, and then we can turn to the class plaintiffs. But we defer to the Court.

THE COURT: This is kind of helpful. It actually brings me to another homework assignment that I have for you all, which is that it's extremely difficult to make sense of the various dockets and with the number of oppositions that have been filed and that have been coming fast and furious, including late last night. In the future, please, when you are filing something that is responsive to another motion or

another request, link to that. Okay. Don't just put the ECF number in the text of the order or the text of what you are naming that document. But make it so that it creates a hyperlink to the motion or letter that you are responding to in the same docket. Now, I understand that you can't cross link, meaning that if it's not in your particular case, you can't link to that docket number, but at least one document in your docket that you are responding to it will make things a lot easier.

The other question I have is -- and this is sort of generally setting the table part -- that for your documents disputes, can they be generally put into buckets? In other words, as I understand it right now, plaintiffs have a number of requests that are seeking particular types or categories of information. Those categories might be responsive to multiple requests from plaintiffs. Is there a way to streamline this by talking about categories of documents that you are seeking and talking about them as what and why and how are they relevant and proportional and talking about the buckets instead of tying them to a particular request or particular dockets, is that possible?

MS. BROOK: We can certainly try, your Honor.

THE COURT: Okay. Yes.

MS. HURST: Your Honor, Ms. Hurst for Microsoft.

I was wondering if suggesting a slightly modified

procedure for future conferences might also help with this issue?

THE COURT: Yes. Hit me with it.

MS. HURST: Your Honor, I would suggest that the parties identify the issues to one another first, meet and confer about those then file what can't be resolved then give the Court the agenda. If we did it in that order, we would have the opportunity to give the Court the docket numbers on the agenda of the matters that the parties were unable to resolve.

MS. BROOK: Your Honor, Ms. Brook on behalf of the New York Times.

I'm happy to do it that way. I guess my confusion is I don't see how that's different from how we have been doing it. I don't want to agree to something if Ms. Hurst and I are talking past each other. What I understand the Court to be asking for is keeping track of RFP 72, 83, and 46 in the Times case, and then RFPs 112 and 33 in the Daily News case and so on and so forth, is obviously not efficient and very confusing. So what the Court is seeking, is perhaps for the next proposed agenda we give to the Court -- we'll always include docket numbers because I understand that's necessary but not sufficient -- and so maybe instead of just grouping it by the New York Times' disputes and the Daily News's disputes and so on and so forth, we do it to issues relating with which

features at Microsoft and OpenAI they are willing to search from or not and include everyone's under that bucket, or for example issues relating to the for-profit transformation of OpenAI and list everything under that bucket and try to give the Court headers based on topic versus just RPFs.

THE COURT: I think you are talking a little bit past each other but I think you are both trying to say the same things just using slightly different words. Yes, I would like to have a coordinated understanding of, here is a general category or an issue, right. And these are the documents or these are the types of documents that we agreed we are going to produce and search for, search for and produce. These are the ones for which there is still a dispute. And if it's helpful then point to the docket numbers.

I was prepared to give you all homework to be done Friday which is to present to me a chart. I really like things in a visual chart manner, and I think this is really amenable to a chart for a chart of headers. And I will say either we can table the document search and production issues for a short period of time for you to produce that chart, and then I can refer back to the briefing that's already happened on each of those issues. Or if you have one or two buckets that are very clear, and you want to talk about them now, we can talk about them now. One or two? Big buckets.

MS. BROOK: We certainly have, your Honor, two big

buckets from New York Times and I believe also Daily News' prospective that we would like to discuss now.

MS. HURST: Same for defendants, your Honor.

MR. NATH: And same for class plaintiffs.

THE COURT: Are your buckets and the plaintiffs' bucket probably the same?

MR. NATH: There are some overlap. We do have a custodian dispute that we think is ripe. I think the Court's order on the New York Times issue is informative for us.

THE COURT: So talk to me briefly about your custodian dispute. Because I would be inclined to order the same homework unless there is a reason that it's different.

MR. NATH: My associate Charlotte Lepic is prepared to argue the details of the custodian. But at least to the custodian we requested, we have five custodians that we requested last time. There are two more at issue. I think the Court's order in the New York Times custodian issue is enough for us and enough guidance for us to move forward on this. So if that's all the Court needs to know, then we don't need to discuss the custodians on our end further.

THE COURT: Let's go from there, and then we will talk about it next time.

MR. NATH: Great.

THE COURT: Do you all want to take a short break to figure out which one or two issues you want to talk about?

MS. BROOK: That would be wonderful, your Honor. We can do it in five minutes.

THE COURT: Five minutes. Great.

(Recess)

THE COURT: Please be seated.

THE DEPUTY CLERK: Court is back in session.

THE COURT: All right. Let's have a bucket.

MR. NATH: Your Honor, we have two issues, one relates to texts.

THE COURT: Texts?

MR. NATH: A collection of texts and social media.

The second is somewhat related to documents but also related to deposition protocol. It's clarification regarding whether we could take a custodial 30(b)(6). We can go in order of those two issues. My partner Jordan Conners will argue the text and social media motion.

THE COURT: Okay.

MS. BROOK: Your Honor, for clarity, those are the class plaintiff's two issues. On behalf of all of the news plaintiffs we came up with two documents issues, and they relate to training data, inspection issues and the like that have come up in our efforts to play in the sandbox.

Number two relates to search terms, though I will preview for the Court that we have a suggested solution that would not require the Court to go search term by search term.

I did want to flag to the Court, it's not documented related, but defendants are objecting to one of our experts. It would be helpful, in terms of moving expert work along, which is happening in tandem with document production, to get a resolution on that today.

THE COURT: Wow. Okay.

MS. BROOK: Thank you.

THE COURT: Go ahead.

MR. CONNORS: Thank you, your Honor. Jordan Conners from Sussman Godfrey for class plaintiffs.

I'm going to be arguing, this is Docket Entry 230, and this is our motion related to text messages and some other electronic messages other than email. The OpenAI response, I believe, is at Docket Entry 243.

THE COURT: 230 and 243. Hold on a second. Let me bring that up.

MR. CONNORS: Your Honor, this bucket, if you will, is about text messages and direct messages specifically from X.com, but it's really a broader issue. It's about electronic messages other than email. All of our document requests ask for these kinds of communications. It's a way we define communications and documents in the request. We have been seeking these since November 2023. We've had extensive meet and confers, dating back -- I think some of the meet and confer traffic that we submitted to your Honor among many reading

materials go back to July. The parties have been talking about this a lot.

There is one major issue OpenAI has taken that has really prevented us from getting anywhere in those discussions. And that's their position that relevant documents in OpenAI's employees' possession and control are not within OpenAI's possession and control. So there are a lot of documents in the case, obviously, that are in central repositories, but there are a lot in the control of the employees. OpenAI has taken the position that to the extent it is a text or a direct message on a social media account or some other electronic message, if the employee has it to maybe on their personal phone as opposed to a phone purchased by OpenAI, that they don't have to search it, and they just don't have control over it, so it's not part of discovery.

THE COURT: That's not -- let me hear from defendants.

MR. CONNORS: Okay.

MR. GOLDBERG: Thank you, your Honor, Nick Goldberg of Keker Van Nest & Peters for OpenAI.

Let me just start by saying that is a wild and drastic mischaracterization of our position.

THE COURT: Okay.

MR. GOLDBERG: Your Honor, there really is no material dispute before the Court with respect to text messages. We've laid this out in our briefing. This is on Docket 234. They've

sought text messages from three senior executives at OpenAI. Those text messages reside on those individuals' personal devices. We read the case law in this district about practical ability. We are very familiar with it, and we have asked those employees for assistance in providing us with that information. I've personally spoken with the counsel for each of those three individuals. That is in process, that is not what our objection is. There is no dispute before your Court about those text messages.

THE COURT: So what is the objection?

MR. GOLDBERG: The objection has to do with X.com messages which is a direct message sent on Twitter. I still say "Twitter."

THE COURT: I say "Twitter" too.

MR. GOLDBERG: But it's X.com, and the problem there, your Honor, is the law is clear that those messages are not within OpenAI's possession, custody, and control. There is a California labor code provision which we've cited in our papers, California Labor Code Section 980. Which explicitly provides that employers can't even ask -- can't even ask employees for their social media platforms. Now, that is clear under the law that we can't --

THE COURT: Under California law, right?

MR. GOLDBERG: Under federal law too, your Honor.

THE COURT: Okay.

MR. GOLDBERG: It's federal law about possession, custody, and control. For example, your Honor, this Sedona Conference looked at this exact issue. It specifically found that no court -- by the way, they've cited none -- that holds where you have a direct message on a social media platform, on an employee's personal social media platform, that that is within the possession, custody, or control of the employer. And you know how we know that, your Honor -- this is a fact they didn't mention in their brief. They told us that they intend to subpoena that same information for the employees. So they tacitly concede that we don't have possession --

THE COURT: Oh, I would not go that far.

MR. GOLDBERG: Fair enough.

THE COURT: There are plenty of times. Have you ever heard a law firm doing a belt and suspenders?

MR. GOLDBERG: I've heard of that.

THE COURT: Just in case we don't get it, we are going to do the slower route of the subpoena. Just in case we don't get consent to take a foreign deposition, we are going to start the Hague process just in case.

MR. GOLDBERG: You are right about that.

My practical point, your Honor, is they are perfectly capable and know how to obtain this information in a way that is compliant with the law.

The second issue, which I actually think is the

threshold issue for your Honor's consideration, it's true, as a matter of law, these X messages are not within our possession, custody, and control.

The more material issue, and I think, frankly, the easier way to resolve this one, is there is absolutely nothing in the record that suggests there is one bit of relevant information in these direct messages on X.

THE COURT: Wait a minute. Who are the employees that we are talking about?

MR. GOLDBERG: For the X issue, your Honor, they've cited four: Chelsea Voss, Nick Turley, Sam Altman, and Greg Brockman. They don't attach to their papers the X messages that they think make these individuals relevant. We do. If you look at Exhibit 2 to our opposition --

THE COURT: That's ECF 34, Exhibit 2.

MR. GOLDBERG: 243.

THE COURT: Okay. Hold on a second.

MR. GOLDBERG: Absolutely. I have it here too, if it's easier for me to hand up.

THE COURT: I've got it. I don't got it.

MR. GOLDBERG: That's okay.

What you'll see when you pull it up, your Honor, is that it's Ms. Voss providing a book recommendation over the X platform. Exhibits 3 to 5 are her re-tweeting other people's tweets without comment. The other examples they cite in their

motion are individuals at OpenAI posting about product releases. There is just nothing in the record that would substantiate their burden, under the Coventry case which we cited, they hold.

THE COURT: Exhibit 2 is a tweet, it's not a direct message.

MR. GOLDBERG: Correct, exactly.

THE COURT: We are talking about direct messages.

MR. GOLDBERG: You're putting your finger on another important issue. There is not even any evidence they've cited that people have communicated by direct message about anything in this case.

THE COURT: I'm barely out of the Stone Age -- well, I was a barely out of the Stone Age user of X when it was Twitter when I was in private practice. I used direct messages. I know how to use them.

MR. GOLDBERG: Whether or not somebody uses them, doesn't mean they used it to communicate about --

THE COURT: Right, but if they are presumptively not public, you are placing a burden on them to show that there were direct messages that were used. Would you like a deposition instead before this happens?

MR. GOLDBERG: No, your Honor, and I don't think that's consistent with the way that the case law works. The case law requires -- it requires that they make an initial

threshold showing of relevance and then there is a proportionality --

THE COURT: Let's take this -- we can take this two ways: One is I read headlines. I don't necessarily read the news. But I seem to recall a published article that quoted Mr. Altman saying something about needing to use copyrighted works to train their large language model. And it might be reasonable to infer that there might be communications before or after with Mr. Altman about that particular thing that he said publicly.

The other way to look at it is I have also been managing a very large case, I believe it was JPMorgan Chase v. Tesla. And in that case, where at least some of the individuals, some of the high-level individuals had less -- well, one of them had less-public statements that related to the subject matter of a lawsuit. I mean, the other one -- I mean, his public statements was the reason for the lawsuit -- that there were messages, and the parties were able to agree that some limited scope -- there is going to be some search of the messages and that some limited amount would be produced. So why is that not instructive here?

MR. GOLDBERG: Well, there are two pieces to that. One, your Honor, I'm not sure what public remarks you are referring to, I'm really truly not.

THE COURT: Going private.

MR. GOLDBERG: The going private remarks -- anyway --

THE COURT: That was the tweet by Elon Musk, JPMorgan Chase v. Tesla.

MR. GOLDBERG: No, I thought you were referring to public statements by Mr. Altman.

THE COURT: My clerk reminds me they were back in January.

MR. GOLDBERG: I do know what tweet from Mr. Musk you are talking about, for sure.

My point, your Honor, is with respect to the relevance of direct messages, there is truly nothing in this record that suggests there is anything in those messages that would be relevant. Even if there is, your Honor, it is not proper under the law, in fact, we would be precluded from asking, under California law, to access those messages from the employees. So there are two fundamental issues here. One, is a relevance issue, and the other is simply just we cannot control those documents. And there are privacy interests involved here that weigh heavily against forcing a company, contrary to a state statute that existing and binds us to demand or ask a fact about their direct messages.

THE COURT: But you just said that you have asked those employees and that search is in process.

MR. GOLDBERG: For text messages.

The California State Statute, Section 980 is limited

to social media, your Honor, so that's the distinction. I agree with you as per text messages, and we had made that ask, and it's underway. With respect to X messages, those are not in our possession, custody, and control. And you've heard my argument about relevance, but I'm detecting a measure of skepticism.

THE COURT: Let me hear a little bit about what and why the California statute matters, the response from the witness.

MR. CONNORS: Thank you, your Honor. Jordan Connors, Sussman Godfrey for the class plaintiffs.

Before we let go of texts, I want to make sure it's clear on the record what's happening. We have been conferring with this issue for a long time. What I heard the OpenAI lawyers say is it's in process. What we hope is that they will search and produce relevant text messages, responsive text messages, just like they would do for email. So I just want to get that clear on the record if that's going to happen or not, and if not, I want to argue about it.

THE COURT: We are not going to argue about it. For that one, I'm going to give you a deadline, one week, report to plaintiffs the status of the search on text messages, what and whether you are producing. My expectation is that since you started the search and you started to ask, that we'll at least get a progress report, if not some understanding of when and

what will be produced.

MR. GOLDBERG: I think that's totally fair, and we can do that.

THE COURT: So one week, that's November 6.

MS. BROOK: Correct.

THE COURT: Thank you.

MR. CONNORS: Thank you, your Honor. And how the text issue was teed up -- still on texts for a moment.

THE COURT: You are kind of winning that one. Go ahead.

MR. CONNORS: It's much broader from when we brought in our motion. We sent a bunch of evidence that some specific employees used text message for business. Maybe they were negotiating a deal with Microsoft or they were talking about training data, they were using text for OpenAI business. So we asked for them to at least search for those custodians. We asked OpenAI itself as part of electronic discovery to ask its own custodians if they use text for work, and if so to include those in their will their search. They have refused to do that and said, sorry, that's work product. We are not going to get into that. I think it's clear that they at least have an obligation to look through their own sources of data to see where there may be relevant information, and we should meet and confer about that, and that was the third ask in our motion. I am prepared to argue about why they have a requirement to do

that. My suspicion is your Honor is already on my side on that one.

THE COURT: You are already winning that one.

I wanted to ask, though, is there any distinction being made between company devices or personal devices?

MR. GOLDBERG: Yes, your Honor. The distinction is that for company devices, to the extent they exist with relevant custodians, those would be obviously within our possession, custody, and control. The text of the three individuals that I mentioned earlier are on personal devices.

THE COURT: What's the company's BYOD policy?

MR. GOLDBERG: The company has a policy. I don't have it in front of me. My understanding is that it says that you generally should not be communicating on your personal device. The law, and this is what we are following for the three individuals, for their personal devices, we go out and ask them for cooperation. And that's exactly what we are doing. We are intending to provide an update on that one week from today.

THE COURT: Okay.

MR. CONNORS: Thank you, your Honor. Let's talk about social media. This was the first of three requests in our motion. Today, OpenAI is arguing that social media is different. If they are communicating about work in a social media, direct message, or some kind of other social media message, they don't have to ask about it because of California

Labor Code Section 980.

THE COURT: They are going beyond "don't have to ask." They are saying that the statute says they can't ask. They are not allowed, by statute, to ask.

MR. CONNORS: Right. I would just say this argument, if you were to accept that argument, it wouldn't be confined to social media. If you look at Section A in that statute, social media is defined broadly. It includes text messages and email. So it's not like we are just talking about X.com direct messages. If you take OpenAI's reading of California Labor Code, Section 980, Subparagraph A, social media is very broad. But that's not what Section 980 means, and this was enacted in 2013 in California. And if this law prevented large corporations from having to ask their employees for discovery into text and email and social media, you would think that they would find one case that limited discovery of California corporations based on Section 980. They haven't done that.

Instead, there are a lot of cases, in this district and in California, that say when you are dealing with federal discovery rules, those rules govern over state privacy or confidentiality codes like this one. OpenAI could have it one of two ways here. Either Section 980 doesn't limit a company's federal discovery obligation, which we think is the best reading of the statute and is a reason why there are no cases saying it limits discovery, or it actually is California

statute that is protecting California corporations from federal discovery obligations in which case it's clearly preempted under the law. So it's either/or, but either way we get discovery into OpenAI should have to ask its employees for relevant documents. Just as, your Honor, if an OpenAI employ went into the office and brought a bunch of folders home and they put work documents in the attic. If we ask in discovery for OpenAI to produce relevance documents, they've got to ask their custodians, where do you keep the relevance documents? Even though OpenAI isn't allowed to go to that employee's attic or ask about what is in the attic, that doesn't shield them from general discovery obligations to find where the relevant materials are. So that's why Section 980 has nothing to do with discovery. It doesn't say it creates some kind of protection to shield that kind of discovery, and there is no case law interpreting it that way.

For those reasons, we think that is just an excuse, and once we get by that excuse there is no other objection. There is not a relevance or burden objection and so OpenAI should turnover those documents, and we should move forward with discovery. Thank you.

THE COURT: Okay. Thanks.

Anything that OpenAI would like to add?

MR. GOLDBERG: Two brief points, your Honor. I would urge your Honor to look at the citations in the Sedona

Conference report we cited which canvases the case law and says, "corporations do not own or control their employees' personal or social media accounts." In fact, "no court has specifically held that, including under the practical ability standard." In this Court they talk about cases which they've cited in their papers addressing state confidentiality laws. Those cases arrive under 1983. They are very, very different from what's going on here. The cases there deal with in the context of issuing a protective order, which you can do all the time. It happens in federal court. And they talk about Section 1980 of the California Labor Code being preemptive. They haven't cited, and I'm not aware of anyways case that has made that holding -- it does not conflict with federal law. What it's saying is that employers can't ask their employees to divulge personal social media accounts.

THE COURT: Is this can't divulge passwords or can't divulge -- it's just a blanket to cover anything related in the social media?

MR. GOLDBERG: It's broad, your Honor. Section B talks about in one section disclosing a user name and password, but in a separate session talking about accessing the account and in another subsection, Subsection 3 divulging any personal social media, so it's all of the above. What it does is it makes it such that we can't ask, therefore we can't control. And if they want it, which, again, I don't think it's relevant.

They've said we made a relevance objection. We have. But if they believe it's relevant, that can be addressed, but the proper way to do it is through subpoenas to these individuals.

THE COURT: Okay. Which have been served?

MR. GOLDBERG: I don't know if they served them or not, but they have been threatening to serve them.

MR. CONNORS: Belt and suspenders, your Honor. We haven't served those subpoenas. We hope we don't have to. We open hope that federal case discovery -- we don't have to go and serve 30 subpoenas on everybody's employee to get their documents and then negotiate we have to negotiate with, I presume, OpenAI's lawyers who would represent all of those employees. We would have a whole lot more discovery issues rather than discuss this in a meet and confer generally for all of them.

THE COURT: Is your issue with the California statute fully briefed? Do you have a response on the Sedona Conference report? I don't have it in front of me right now, and I will probably need to take be a little bit of time to digest it and figure this out. I'm certainly leaning plaintiff's way, but I want to give you a chance to talk about if there is more that you wanted to say about it.

MR. CONNORS: Yes. So we've cited a series of cases in our brief that generally go to the proposition that an employer has control over an employees' documents. One of the

cases talks about an employee's gmail account, for example, which I think is very analogous here. There wasn't a case cited on either side that talks to X.com direct messages. But I don't see how that would be any different. Certainly, under Section 980, which includes texts and email under what is defined as social media, there shouldn't be any difference. I think we do have persuasive cases in our brief that refute that one sentence from Sedona papers.

I do just want to respond, which I didn't do earlier, I should have, to the relevance point. I want to read just two of these X.com public posts that they didn't attach. We linked to them in our brief. Perhaps we should have copy and pasted them. I'll read a couple for your Honor so you can see how clear this is.

THE COURT: Links are actually fine.

MR. CONNORS: If you follow the link you will see this X post from Sam Altman. He says, "everything we release, ChatGPT 3, DALL-E, Codex is just a mile marker, and we have a roadmap to make them 10X better. Please come help us do it. DMs open." DMs open, so that obviously means direct messages. He is soliciting direct messages about recruiting for these products that are very much at issue and accused of being part of this copyright infringement case.

Another one that they didn't mention, this is from Greg Brockman. He says, "anybody who would like to build an

app on our API, please ping me. May be able to accelerate an invite to you." These are high-level OpenAI employees who were inviting DMs, inviting pings on X.com to speak with people about these products. They are using it just in a way more traditionally some people use email. They are using it for work.

THE COURT: They're using it for work.

MR. CONNORS: Absolutely. So we've established that and a California labor code is not going to prevent a federal court from getting broad discovery to get to the truth of the matter. We don't believe that should be the ruling, and there is no case saying that it is.

THE COURT: Okay.

MR. CONNORS: Thank you.

THE COURT: Anything else you want to say? If they're affirmatively using their X accounts for work and soliciting private communications over a device for work purposes, I don't know how this is different from the discovery that was agreed to by the parties in *JPMorgan Chase v. Tesla*.

MR. GOLDBERG: I hear your Honor's comments. I don't think that saying to somebody, "come work at OpenAI on products and DM me about that," is relevant to this copyright case. I just don't. There has got to be --

THE COURT: What about Altman's messages about this is just a mile marker for our big plan.

MR. GOLDBERG: Yeah, come work here. I truly don't think that has anything to do with the issues in this case. And we are seeing discovery in this case spiral. I think this is, quite frankly, a little bit of a detour and there are much more core issues in this case.

THE COURT: I don't think it's a detour. When you have heads of the company who have made public statements about issues -- about issues that are directly being litigated in this case, I think it's a problem.

MR. GOLDBERG: I hear what you are saying, your Honor. I don't think that "DM me" falls into that category. But regardless, your Honor, there are two issues. The second one has to do with possession, custody, and control. We just don't have it.

THE COURT: Have you briefed it in the form of seeking a protective order because that's really what it sounds like you are asking for?

MR. GOLDBERG: It's briefed consistent with a letter motion in accordance with your Honor's guidelines. If you would like a protective order, we can do that too.

THE COURT: You mean like a proposed protective order? I meant, is there anything else you want to add that you haven't put in you are arguments?

MR. GOLDBERG: No, I do think it's been briefed, yes.

THE COURT: Is there anything else you want to add?

MR. CONNORS: Yes, your Honor. The only thing I would like to add is, as you are considering this and considering issues an order on this, the way the parties have been proceeding so far, just by necessity, is we have looked through public documents, it's not always easy, but looking at X.com public posts or other documents in the production to try to prove that particular employees -- look, that one uses X.com for work. That one uses texts for work, and then asking OpenAI, hey, would you produce that employ's documents? We think that's backwards. We think they have an obligation --

THE COURT: I'm not setting this as a new standard for showing relevance, no. No. I understand that this is part of a robust meet and confer. It's similar to in the old days when there were paper documents or emails were produced on paper and they said, who is this person copied here? They seem to have replied with something of substance, and you talk to the other side and you say, who this is this person? Can we have their documents? Who are they? What do they do? And then you have that discussion and maybe it ripens into a motion to compel or a motion for a protective order. But I don't see this as being really materially different.

MR. CONNORS: Right. But what we would ask, because we think this is the same, is for the custodians -- OpenAI or Microsoft has agreed to -- that they just simply ask those custodians what they use for work, what they use to communicate

for work, maybe it's text, maybe it's X.com, maybe it's some other electronic message platform. And they let us know. They don't have to agree to produce it, but at least we know this is what they use for work, and we can make an argument about proportionality and relevance once we have that information. But the burden shouldn't be on us to try to find evidence of what their employees use to communicate for work just like any other kind of discovery.

THE COURT: Okay. I almost never used Rule 33 when I was in practice. It seemed easier to go to document requests or depositions. Would Rule 33 interrogatories, particularly tailored interrogatories for that reason be helpful? And I would not limit you to 25 under the rule. And I would not limit you to future interrogatories, if necessary. You might need to make a minimal showing. But it seems like if what you are asking for is as simple as, what are the platforms or media or avenues that each particular employ uses to communicate for work or have you ever used for work purposes, that would probably help. Is that what you are asking for?

MR. CONNORS: I think that's a great idea. We think this is general meet-and-confer information that, under the Sedona Conference principles, the parties exchange. They haven't done that. I think under their current position in the case they would resist such an interrogatory. But I think that's a very efficient way to proceed and as long as they will

answer them, we would be happy to serve them. We would like the information any way we could get it, as quickly as possible because it's sort of fundamental to the next step in written discovery.

THE COURT: Why don't you do this, serve those interrogatories, serve those subpoenas. And the parties have one more week to meet and confer on whether this motion is going to be withdrawn. Okay. In other words, I think you all see which way the wind is blowing. If they are not going to get it through agreement, we will take it the slightly longer way. If we take it the slightly longer way, I will look at costs under 37(a)(5). I'm not saying I will apportion them. But I will look at -- if we it have to take the long way what are the costs.

MR. CONNORS: Serving interrogatories on defendants is one thing. I think that will get the information quickly and move us along.

You want us also to serve subpoenas on all the employees for all of their --

THE COURT: The belt and suspenders.

MR. CONNORS: I think that may take very long timing.

THE COURT: Why?

THE DEFENDANT: We'd have to seek discovery from each individual employee and not the company as whole. They are going to have their own objections. I don't know who is going

to represent them, maybe it would be the same lawyers. But that seems incredibly expensive and inefficient.

THE COURT: It might not be that expensive for you. That's why I'm cautioning the defendants.

MR. CONNORS: Okay. Thank you, your Honor.

THE COURT: You may agree to hold the subpoenas in abeyance for a short period of time to see if you can work this out at lower costs. I go back to my initial guidelines in the beginning of this conference.

MR. CONNORS: Okay. And I suppose for custodians who have already been agreed, we hope their documents have been preserved. That's something we can discuss in the meet and confer, I suppose.

THE COURT: Yes.

MR. CONNORS: Thank you, your Honor.

MR. GOLDBERG: One final point, to the extent that your Honor is looking at this California labor code provision, and they made this preemption argument that it's preempted by federal law, that's something we addressed briefly in our motion. I think that's a significant issue. If your Honor is inclined to reach that issue --

THE COURT: I asked you if there was anything else you wanted to add, and you said you were done, and it's in your papers, so I think you're done. I can also do research. I really don't need 100 filings in a month to get through this.

This is not rocket science.

MR. GOLDBERG: Okay.

THE COURT: Next issue.

MS. HURST: Your Honor, the defendants also have issues. Mindful of the Court's time, I'm wondering if we could alternate?

THE COURT: What are your issues?

MS. HURST: Microsoft would like to raise its prompts and outputs motion, which is Docket Number 284 in the New York Times case. Okay.

THE COURT: Okay.

MS. HURST: That's our only issue.

MS. YBARRA: Your Honor, OpenAI has two issues related to plaintiff's production. One is Docket 276 in the New York Times case regarding the Time's production of damages or documents related to damages and fair use. And the second is Docket Number 232 an Authors Guild case, also relating to damages documents.

THE COURT: Are we talking about damages now?

MS. YBARRA: Because we're in fact discovery, your Honor.

THE COURT: No, that's not good enough. That's going to have to wait.

What about the fair use? Doesn't that relate to the production issue that I had said at the last conference that I

hoped to rule on by today, and haven't ruled on yet?

MS. YBARRA: I believe there is some overlap, your Honor. I don't think it's entirely duplicative. My partner Andrew Dawson is going to address that one.

THE COURT: Let's put that one on hold also just because I have made progress on that. I just haven't issued an order yet. I would like to wait on that one because I'm hopeful that the order that I issue will help streamline some of the issues a little bit.

MS. YBARRA: Understood, your Honor.

THE COURT: Prompts and outputs, didn't we talk about this last time too?

MS. HURST: Your Honor, you talked about it in the context of the copyright claims. However, our motion is directed to the dilution and the common law misappropriation claim, where the prompts and outputs are disclosed in the complaint, and they serve, in our view, a different role, a very different role than the ones that were at issue on the copyright claims in the prior discussion.

THE COURT: Can you just briefly describe that issue for me.

MS. HURST: Yes, your Honor.

So using a dilution claim as the lens here, the New York Times is claiming that they were able to cause Bing Chat, let's just say for purposes of our discussion, to

mischaracterize the contents of a New York Times article or to inaccurately summarize it or to give a broken link, things along those lines, your Honor.

And then they claim that the substance of that interaction between the prompt and the output reflects a consumer state of mind that will tarnish the New York Times because the consumer will attribute the error or mischaracterization to the New York Times and that it's Mr. Altman who caused that tarnishment to occur, your Honor. It's not possible to understand that theory of the claim without looking at both the prompts and the outputs. Indeed, the Times recognizes this because they disclose substantial portions of the prompts and outputs in their allegations. Your Honor, it's our position that they can't do that to come into court and then disavow all that and say they won't rely on any of it for purposes of their case because then there is no claim. In fact, they weren't willing to disavow reliance on it entirely. They said, no, we are relying on it for purposes of Rule 12, and Rule 26. And we will give the allegations in our complaint to our experts.

Your Honor, it's our view that under 502(a) and the cases that we cited, that requires disclosure in these circumstances. Both because there has already been a personal disclosure and because they put this at issue in a way that extends to their work that's underlying facts, your Honor.

These are facts, the contents of the prompts and outputs that are alleged in the complaint in connection with these other claims.

THE COURT: Okay. Let me hear briefly what the objection is.

MS. CRONIN: Good morning, your Honor. This is Emily Cronin from Sussman Godfrey for the Times.

At a high level, we disagree that this is different than the motion that OpenAI brought and that your Honor ruled on last month. We think that the issues overlap and, we think the request for production to overlap, considerably. That motion dealt with Exhibit J, but it also department with all the prompts and outputs cited in the complaint. Just like with Exhibit J, the Times is not intending to rely on these prompts or outputs in summary judgment or at trial either.

Just to clarify a point, the Times' experts are not planning to rely on these either. All we said is that in so far as we give our experts a copy of the complaint, the publicly filed complaint that this Court and defendants have access to, they will read those allegations and see the examples cited in the complaint. Our experts are not relying on the underlying work product or any of the Time's counsel's investigation in their analysis in this case.

And we disagree with the basic premise that you can't assess the merits of the misappropriation claim or the

trademark dilution claim without looking at this attorney work product and privileged communications. Our position is that the evidence of this will come from defendant's own data, so Microsoft's output logs, which with they've agreed to produce. That will show what user prompts go into certain queries and what outputs are generated.

And the key to these claims is that Microsoft products, let's say Bing Chat, are reproducing or summarizing Times and Wirecutter content in a way that is unfair and that is misappropriating the Times' works. And in the case of the trademark dilution claim, producing hallucinations. So the attorney investigation and the documents and communications about that attorney investigation are not probative on these claims and again, the Times is not going to rely on these investigations or the communications about this investigation in its case in chief at trial.

If your Honor has no further questions, I'm happy to rest on what we briefed in the papers.

THE COURT: Okay. For Microsoft, do you want to add anything other than what's been in your papers? I'm not ruling from the bench on this.

MS. HURST: Thank you, your Honor.

I would say it's not the rule that you can plead facts and then disavow them after the litigation has started to proceed. And let me just say the practical problems with that

here. Your Honor, in focusing again on the dilution claim, we don't understand -- this is supposed to be about consumer mental impressions. Without those prompts and outputs, there is literally nothing to tell us what this claim is about. And we cannot wait until we get some expert report a couple of months before the summary judgment deadline to understand this claim. The prompts and outputs are necessary to understand what state of mind and what actions are really at issue. Is it broken links? We can address broken links and whether that's properly the basis for some kind a tarnishment claim. Is it giving an outdated piece of health information?

Your Honor, them giving this to their expert, whether the expert says they rely on it or not, is a problem. And let me give the Court an example of why. What if they used an automated approach and they pinged the models with hundreds or thousands of requests to come up with a prompting strategy that would enable them to generate these allegedly false things and bad things attributed to the Times? And then their expert looks at their complaint and the expert is able to get the benefit of that, start a new with the benefit of that and then never be cross-examined on how hard it was to actually come up with that approach. And in that's like saying, the incidents at which this occurs is in reality, one-fifth of 1 percent, but we don't know that because all of the work it took them to get to the subject matter that their expert now gets a head start

on can't be cross-examined, your Honor, and that's just fundamentally unfair.

THE COURT: I have two questions. One, I will ask the question that might take some digging from the other attorneys who aren't speaking. We talked about this at the last conference. It's 134-page transcript. If somebody wants to point me to the pages roughly where we talked about it -- you have it already.

MS. CRONIN: Your Honor, page 63 of the hearing transcript is where OpenAI's counsel made a lot of the same arguments.

THE COURT: I'm going to just put a tab on that right now. Because the next question I'm going to ask is, you made a point -- what's your name again?

MS. CRONIN: Emily Cronin.

THE COURT: You made a point, Ms. Cronin, about you're going to be getting information from Microsoft's own output logs. Explain to me what that means and what the information from Microsoft's output logs, which I'm guessing is going to be produced in discovery?

MS. HURST: It's actually quite complicated, your Honor. That's 15 petabytes of data.

THE COURT: I don't even know how many exponents that is.

MS. HURST: Exactly, your Honor.

Whether that data ultimately shows anything at all relevant this to this question is highly uncertain and will be the product of a long chain of events.

THE COURT: Talk to me about what the Times would expect or hope to get from the Microsoft output logs. And then let me hear a little bit more background. This is educating myself on what the output logs are. What I'm also hearing is this might be premature, depending on what is produced when you say you are producing output logs.

MS. CRONIN: What we are hoping the output logs will show or what we expect them to show is exactly what Microsoft says it it's from this work product and from the Times' investigation and complaint, which is how these outputs are generated. What user prompts go into their Bing Chat, and then how --

THE COURT: I'm sorry to interrupt you, but these would presumably be non-New York Times users, just other anonymous users putting in prompts and then what the output is from those prompts?

MS. CRONIN: Yes. And to point to the misappropriation claim, this deals with the Hot News Doctrine. How are Microsoft products, how is Bing Chat retrieving recent articles or Wirecutter recommendations and providing them, either in summary form or verbatim, to its users. Which, again, the high-level issues that I'm hearing, which is on page

63 of the hearing transcript, is that Microsoft needs to know how many and what kinds of prompts users need to enter in order to generate misappropriating or hallucinatory content and how hard it is for a user to generate that. In addition to the output logs, it seems feasible that they could test their products themselves, and their experts would test how many times you have to ask Bing Chat before it produces a hallucination of a Times article and falsely represents that it's the Times or how hard it is to get them to reproduce yesterday's Wirecutter article on Christmas gifts.

THE COURT: Okay. Go ahead.

MS. HURST: Your Honor, if we ask Copilot or Bing Chat today "what's the contents of a New York Times article," they will say that's behind a paywall. I can't give it to you. This is part of a problem. It's not like we haven't tried to figure out how they got these results and duplicate them. We have.

And your Honor, if you think about the difference between this and Exhibit J, in connection with the copyright claim, Exhibit J was for them to enable to say, we have a good faith basis for inquiring whether this is in the training data. That's like a binary thing. There are some complications related to it. But there is another clear source of data, it's either in the training data or it's not. That's not what we are talking about here. Here is the actual contents of the

prompts and outputs that are at issue. It's not about whether there was something in the training data. It's about the actual behavior between a user and a model to produce an output. And their claim, their dilution and misappropriation claim rely upon that behavior, the only facts of which are the allegations in the first amended complaint, which they acknowledge had to be disclosed in order to state the claim because they summarized them. They disclose and summarize them in those two claims to meet Rule 12.

On the data, your Honor, so here is where we are, there are 15 petabytes of this user session data. We've primarily meeting and conferring with Ms. Maisel about this. She has taken the lead for the plaintiffs on this issue. We've given them samples of the data and the fields listing to enable them to formulate a sampling methodology. It may be that there will be further disclosures needed to be made in aid of their ability to do that, your Honor. But in order for them to get at this issue, they are going to have to do a sample of some kind. We are going to have to then see if we could query the database to do that. Nobody knows what volume that is going to look like, and nobody knows what's in there. This isn't like the training data where you load it up on a machine and go look at it. This is total speculation whether anything in there would support a theory that consumers are confused about the operation of this model with respect to the New York Times or

not.

And so, your Honor, it can't be that they plead themselves into court to take discovery of a 15petabyte trove of data in a total fishing expedition without having to produce the facts that form the basis for their allegations.

THE COURT: Talk to me a little bit more about the output logs. Is it that there are logs maintained of any time anybody queries the database and then you are trying to get at a subset of that?

MS. HURST: Your Honor, there is numerous fields or items of data that are collected in connection with each session between the user and the service.

THE COURT: In this instance we are talking about a particular user or are talking about --

MS. HURST: No.

THE COURT: So it's not limited to, say, a New York Times' affiliated user, it could be users at large?

MS. HURST: Users at a large, your Honor.

THE COURT: What I understand from what you are describing is that it's taking a long time because you are trying to do a sample. Is there any dispute about how long it's taking or what the process by which you are getting these output logs?

MS. HURST: Not at the present time, your Honor.

THE COURT: Okay. Well, in that instance then, I'm

going to deny the request on grounds of privilege and confidentiality reasons without prejudice to renewal because I think this will depend on the progress of what you are getting from the output logs. What I'm hearing from the New York Times and the press plaintiffs is that, at least now, I don't think it's been shown that you need this information if you are able to get essentially sanitized or clean non-privileged information through this process. Because what I'm hearing from the New York Times is that they ran these certain queries to show that they had a good faith basis to plead that the search engine was producing either broken links, misinformation, false information, *et cetera*, right. But that is not limited to, I think what they're claiming at large, which is the interactions with this AI process turns them out. I don't think they are saying that we are the only ones who've gotten broken links or misinformation. So that's why I think that the sampling of the output logs, onerous and burdensome and incredibly data heavy as it may be, may be a way to avoid having to get to the privilege and confidential material.

(Continued on next page)

MS. HURST: Your Honor, assuredly they are not saying we are the only ones limited to this. We are the ones saying they are the only ones limited to this, that they have artificially produced these results.

Your Honor, I believe we are entitled to examine that contention when they have disclosed the facts of the prompts and the outputs in their pleading allegations.

THE COURT: That's my ruling right now. It's denied without prejudice to renewal. We're going to have to take it up at a later time when you're going to have to show me that the output logs are not sufficient.

All right. And this was ECF 284. Was that?

MS. HURST: Yes, your Honor.

THE COURT: All right.

MS. BROOK: Your Honor, Davida Brook, Susman Godfrey. If it would be helpful to play MC, there are four issues remaining of all of the issues that the parties decided to raise during the break. Do you want me to give you the topic header of what they are?

THE COURT: Yeah, let's see what they are so I can see how we can move this along.

MS. BROOK: Thank you, your Honor. So from the class plaintiffs, their second issue was whether or not they can take a custodial 30(b)(6).

THE COURT: But I thought, and I'll let Mr. Nath let

me know, but I thought you had sat that given my earlier discussion and ruling on the custodians issue in the other case that you maybe had enough guidance to go forward on that.

MR. NATH: That's true as to our motion to compel additional custodians. We had briefed as part of the deposition protocol whether we can take an early 30(b)(6) on a custodian of records issue. We've met and conferred about it with the defendants several times, and I'm mindful of what your Honor said earlier today and the reasons you gave for why an early 30(b)(6) makes sense in this case are spot on.

But I just wanted to let the Court know that we have met and conferred a lot about this and have really reached an impasse on the question. So it would be very helpful to have a ruling now, in particular because I'm hoping we can have a conference in December and I'm hoping --

THE COURT: Oh, you'll have a conference in December.

MR. NATH: And I'm hoping to take that deposition before then so we can streamline some of these discovery issues.

THE COURT: Let me hear from -- is this directed towards OpenAI or Microsoft or both?

MR. NATH: Both, your Honor.

THE COURT: Let me hear from defendants.

MS. HURST: Your Honor, for us there isn't any dispute over custodians between Microsoft and the class plaintiffs.

We've produced from 13 custodians. There's two more under discussion. We don't even have a dispute. So I don't see how there would be a basis for any deposition about custodial discovery.

THE COURT: Let's hear from OpenAI.

MR. GOLDBERG: Nick Goldberg, Keker Van Nest & Peters.

Two issues, your Honor. One is the scope. We heard your Honor's guidance this morning. The scope that they've laid out in the briefing and in our meet and confer discussions for custodial topics is exceedingly broad. And my practical concern is that it's going to lead to protracted disputes in the future about whether they're taking merits depositions now, seeking testimony on merits issues now. For example, they seek to include as custodial topics things related to the storage and retention of all information that can be the issues directly related to the training models, the employee use of software programs or tools, the structure of relationship between the defendants, and other topics not directly related to the merits of the case, which haven't been disclosed to us.

THE COURT: Okay. Stop. By Friday, meet and confer. By Friday, one joint letter. One joint letter on what if anything you have been able to resolve about an early 30(b)(6) regarding custodial issues.

MR. GOLDBERG: Yes. Yes.

THE COURT: I understand you're referring to things

you filed before this conference. And I opened the conference with some general guiding principles. Go back. Meet and confer. See if there's a limited 30(b)(6) that you can agree to.

MR. GOLDBERG: Yes, your Honor. With that guidance is helpful, we can do that. And then the second --

THE COURT: And then if there remains a dispute about this issue, joint letter. And I understand with this case that joint letter might be you each have a page and a half of your positions, rather than writing actually a joint letter. That's okay. But no more than three pages, one joint letter on this issue by Friday. Okay.

MR. GOLDBERG: Understood. And your Honor's guidance this morning, which we heard, is that if there are custodial depositions, it will be mutual. So we will take that under advisement. Thank you.

THE COURT: Okay.

MS. BROOK: Your Honor, Davida Brook, Susman Godfrey on behalf of New York Times. If I can extend that homework assignment to us as well, we would also like a custodial 30(b)(6). So rather than doing serially, we will also participate in the joint letter.

THE COURT: Yes. Okay.

MS. BROOK: Three left. In no particular order for the Court's consideration, there is the training data sandbox

issues, your Honor. Then there are the search terms ones, which again, we have a proposal that does not require the Court go search term by search term. So that one, dare I say, may be more efficient. And then there is the issue of their objecting to one of our experts.

THE COURT: I think I'm going to take up the search term proposals only. But before we get into that, can you point me to, at least a few of the beginning ECF numbers for the other two, the training data sandbox issues and the objection to the expert.

MS. BROOK: Yes, I can. I will let Ms. Maisel speak to the training data ones.

MS. MAISEL: We don't have ECF numbers for the training data. This is Jennifer Maisel from Rothwell Fig. These issues have recently arisen over the past couple weeks after we started our inspection, and I personally went to the sandbox and can speak to those issues. If your Honor requires briefing on this, we can do that but we have one basic request there.

THE COURT: So they haven't been raised in a filing yet?

MS. MAISEL: No filing. We raised it in the joint agenda filed.

THE COURT: Okay. We're not going to do that one today. Yes?

MS. BROOK: Per the expert issue, it's on the Daily News docket, your Honor, and it's ECF 163, 166, and 167.

THE COURT: 166 and 167. Okay. I'll have to take a look at that after the conference.

For the training data issue, why don't you try to see if you can file a joint letter on what that issue is by Friday.

MS. MAISEL: Okay.

THE COURT: Again, three pages. It can be one and a half pages each if you can't actually formulate a joint letter.

MS. MAISEL: Thank you, your Honor.

THE COURT: All right. Let's hear about the search terms issue. And point me to the ECF number and the case.

MR. FRAWLEY: Yes, your Honor. Alexander Frawley for The New York Times.

This is Dkt. 228 in The New York Times case, and this was a search term motion filed against OpenAI. And their opposition brief is Dkt. 233 in The New York Times case. And just as a quick recap, we talked about this at the last conference briefly, and the plan of action was that OpenAI would make a counterproposal to us, and that we would then meet and confer. And that's happened. There's been five proposals exchanged back and forth.

What we're seeing now is that the parties aren't disagreeing about the relevance of particular terms. The entire disagreement, and it's really a gating issue, is the

total number of hit counts. So as you might expect, our terms are broader. Their terms are narrower.

So rather than renew our motion and ask you to go through a spreadsheet and say yes to term one and no to term two and so on, we think a more efficient proposal would be for your Honor to provide guidance on the total number of hit counts where this search term dispute should end up. We're proposing somewhere between 550,000 and 600,000. OpenAI has told us that they think it's more like 240,000. And if your Honor could provide that guidance, then I think the parties, I'm confident the parties can resolve the issue without coming back to you again. And really the onus would be on us because we would have a hit count we're working with and we would have to be selective and careful about what terms we chose.

THE COURT: Okay. So run that by me again. The hit count that you're proposing is?

MR. FRAWLEY: What we said in the agenda was someone between 550,000 and 600,000. And I just want to make one quick important clarification. This applies to both The Times case and the Daily News case. We've been making proposals on behalf of both parties. So I think that context is important for the hit counts.

THE COURT: Okay. And the Center for Investigative Reporting case just got consolidated, right?

MR. TOPIC: Matt Topic for CIR. Yes, we were just

consolidated into this, correct.

THE COURT: Welcome to the party.

MR. TOPIC: Fun to come to the microphone.

THE COURT: So have you been engaging in any of this discussion about hit counts?

MR. TOPIC: We have not. I suspect we are not going to rock this boat other than on issues maybe that are really specific to us and really important. So I'll defer to the other plaintiffs' lawyers.

THE COURT: Okay. And, Mr. Frawley, come back up. And OpenAI wants a lower hit count; is that right?

MR. FRAWLEY: Yes, that's right. They proposed to us, I think by e-mail, they think an appropriate number is 240,000 for the top line hit count.

THE COURT: Okay. We did this before.

Mr. Gratz, you will get a chance to speak. I promise.

But just roll back a little bit. When you're saying we're talking about a total number of hit counts or a top line hit count, what does that mean? Walk me through that.

MR. FRAWLEY: The parties are exchanging search term proposals. There's dozens of search terms in each party's proposal. Each time OpenAI has provided hit counts, both on a term by term basis as well as a total hit count basis, and that's for all 26 currently agreed upon custodians.

So when we send them a spreadsheet and we say here's

our search term proposal, they send us the spreadsheet back and they tell us "your proposal will result in us having to review, the most recent one, 632,000 documents for the 26 agreed upon custodians." And that's why -- that's where we came up with our number of 550 to 600,000 because we think we're pretty close. We see some terms that we should refine in our proposal, but we don't think we need to refine too many to get to what's a reasonable amount of documents for this case, especially for these cases. Sorry.

THE COURT: Pause. Does hit count correspond or correlate to number of documents?

MR. FRAWLEY: Yes, it does. And my colleagues wanted me to clarify, it's a de-duplicated amount. So they've already gone through the process of figuring out -- obviously some search terms will hit on the same document. So they've taken that into account when they say the total hit counts is 632,000 for our most recent proposal.

THE COURT: I see. And each hit is a document?

MR. FRAWLEY: Yes, your Honor.

THE COURT: Okay. And your thought now when you're going forward with your search term proposals is that whatever top line hit count you get told that you're going to have to work with, that will enable you to better formulate and choose which search terms you might want to use and how to maybe streamline.

MR. FRAWLEY: Exactly, your Honor.

THE COURT: All right. Now I'm going to ask Mr. Gratz to stand up again and talk to me about what the difference means for your clients in terms of 240,000 versus 600,000 or 500,000.

MR. GRATZ: Right. And the difference, your Honor, is that the current -- their current Sortium proposal for the current custodians, hits around the number they're asking for, within a few tens of thousands. And their current Sortium proposal is a little flabby. There are a bunch of terms, that I'm happy to go into one in particular, where they're not really paring this down as much as they could. And we feel like we in our proposals are paring it down to really what's necessary and important. And getting them hundreds of thousands of documents.

THE COURT: Yeah. But, again, talk to me practically speaking and logistically speaking, what's the difference between 240 and 600?

MR. GRATZ: Right. And obviously, your Honor, the difference between 240 and 600 is the difference -- that number of documents to review and produce. And what does that translate into? That translates into dollars and that translates into weeks.

THE COURT: Yep. Now pare it down more. How long does it take? I'm assuming, but maybe I shouldn't assume, that

you've done some sort of projection for either time or attorney count. So talk to me about that.

MR. GRATZ: So with our -- with the machine that we have running, an additional 600,000 documents, is about 100 days of reviewer time. Can we do that? Sure.

THE COURT: Wait. 100 days of reviewer time. How many reviewers? Are the reviewers presumed to be reviewing 8 hours a day, 12 hours a day, or 24 hours a day?

MR. GRATZ: Reviewers not presumed to be reviewing 24 hours a day.

THE COURT: I'm not saying your making the reviewers review 24 hours a day, but when you say 100 days, does that mean 2,400 hours?

MR. GRATZ: That does not mean 2,400 hours.

THE COURT: So talk to me in terms of hours and number of reviewers.

MR. GRATZ: Ten hours a day, forty reviewers is the basic --

THE COURT: Forty reviewers.

MR. GRATZ: -- basis for our estimate.

THE COURT: So ten hours a day. That's 40,000 hours.

MR. GRATZ: I think that math is right, your Honor.

THE COURT: So that's 40,000 hours to -- that's 600,000 hits, to finish 600,000 hits; is that right?

MR. GRATZ: I think that's right, your Honor.

THE COURT: So it would be 20,000 hours to get to about half that.

MR. GRATZ: Correct.

THE COURT: Okay.

THE DEPUTY CLERK: If your cell phones are on, just please silence them.

THE COURT: Sorry?

THE DEPUTY CLERK: Someone's cell phone just went off.

THE COURT: So if you're talking about each reviewer is doing a 200-hour month, this is where I think to myself I'm really glad I'm not an associate in a law firm anymore.

MR. GRATZ: Indeed, your Honor. With apologies to my associates and the contract reviewers that are addressing this stuff.

THE COURT: Right.

MR. GRATZ: And this is, you know, we think that getting to a number that is quite materially smaller than 600,000 still gets the plaintiffs what they need on these, the sort of concrete issues that are at the core of this case. And we are letting them choose how to narrow down the documents. And we think there are lots of good ways to do that that we have proposed and that get them what they need.

THE COURT: Mr. Frawley, you're like popping up and down, so come on up.

MR. FRAWLEY: I'm sorry. Thank you, your Honor.

THE COURT: No need to apologize.

MR. FRAWLEY: Alexander Frawley again for The New York Times.

Your Honor, we think that 600,000 is appropriate. It's a big case. It just got bigger today with the CIR being added. So there are already millions of works at issue. We think it's proportional to the needs of these news cases. And in complicated, important, technical cases like these, I think I'm aware of cases where there's been over a million documents produced. Even our proposal for 600,000 documents to be reviewed is going to result in fewer being actually produced because of responsiveness and privilege calls.

So, at the end of the day, if we end up with a couple hundred thousand documents out of our proposal, we think that's more than reasonable and proportional to the needs of these cases.

THE COURT: Go ahead, Mr. Gratz. Do you have anything to add?

MR. GRATZ: The only thing I would add is, to the extent that there was -- to the extent there is an increase over what we proposed and basically our set of search terms, we actually, I mean, we really also need to make sure this includes any additional custodians. Right? This is it. This is the number that we are going to review. That, for example, that the additional, the addition of further custodians does

not require sort of further increase. That we're talking about like the number that we are all aiming.

THE COURT: I'm actually more inclined to think that if there's added custodians, that you'll use the same search terms. No?

MR. FRAWLEY: Yes, your Honor.

THE COURT: Run the same search terms.

MR. GRATZ: Right. And what we're saying is like, as we were discussing before, we're going to do some hits on proposed custodians. We want to like to the -- we want this to be the all in number. Like I agree that it's going to be the same search terms for those additional custodians.

THE COURT: Okay. We may be cutting it too finely at this point. If you really want me to pick a number, I'm happy to pick a number, but let me go through some of this.

If you're talking about 200 reviewer hours per reviewer per month, then if you had 100 reviewers, if you had 100 reviewers, you could get 300,000 hits reviewed in a month.

MR. GRATZ: I think that math is right, your Honor.

THE COURT: Okay. And you would be anticipating rolling production anyway, right?

MR. GRATZ: Yes, your Honor.

THE COURT: Okay. If you were to be doing a rolling production -- this is exactly the kind of micromanaging I told myself I wasn't going to do. But I'm going to ask the

questions anyway just because I can't help myself.

If you were to be doing a rolling production, have you talked about whether the rolling production would be based on terms or custodians or something else?

MR. GRATZ: No.

THE COURT: Okay.

MR. GRATZ: We haven't made any agreements about like what order we're reviewing everything in.

THE COURT: Right. And I know Ms. Brook is trying to make this very, very streamlined. And I'm putting my finger in there and making it worse. What's that?

MR. GRATZ: And I am trying not to micromanage the document production machine to --

THE COURT: Right. Well, I guess because you raise the issue of some of the search terms might be a little flabby. Maybe they are. And I feel like I'm really just picking a number out of a hat right now without really understanding how it's going to be produced.

But here's what I'm going to do. I'm going to give you 500,000 search terms as a cap. But I want you to meet and confer about a rolling production that is doable, that gets the plaintiffs what they think are the most important or the hotter docs first. Okay? Or just to get, you know, to make it easy, easier to get documents produced, and get that started.

And then you should be continuing to talk about what

you're getting in this rolling production. Because it may get to a point where if you've produced the equivalent of like 250,000 documents, that they may say, all right, yeah, yeah, we got enough, or we can turn down the spigot a little bit, or why don't we wind it down a little bit. Right? So I'm going to impose right now a ceiling of 500,000 hits. With guidance. I mean, none of you, and particularly the recently barred attorneys who are probably not in this room, want to be doing more doc review than is necessary.

MR. GRATZ: Absolutely, your Honor. And the only thing I would add is to that is the flabby search terms don't actually result in more documents getting produced. They just result in pulling in nonresponsive documents that need to be reviewed, but then which then don't end up getting produced.

THE COURT: Right. But that's the cap.

MR. GRATZ: Yeah, the 500.

THE COURT: So work on the flabby search terms, and you can get down to 500,000 maybe, and then work on coming up with a rolling production that, you know, you may be able to check in at periotic benchmarks during the rolling production to say are we still full bore on this, or are there things that we can start putting on the back burner because of what we're seeing.

MR. GRATZ: I will remain hopeful that there will be a point at which the plaintiffs will say "I do not wish for

anymore documents."

THE COURT: Well, they may not say I don't want anymore but they might say, you know, let's focus on this search term over another one.

MR. GRATZ: Sure.

THE COURT: This set of hits over another one. And let's let that be an iterative process that doesn't involve me.

MR. GRATZ: Absolutely, your Honor.

THE COURT: Where you can get, as the Rolling Stones say, you can't always get what you want. But my hope is that you'll get what you need on both sides.

Mr. Frawley, you had something else you wanted to say?

MR. FRAWLEY: Yes, just very briefly. That process makes sense to us. Thank you, your Honor.

And I think one reason it makes sense is because, for example, as we get a document that may cause a new potential search term to come to light, and then we might, as part of this process -- no, hear me out.

THE COURT: Good thing facial expressions aren't reflected on the transcript. Go ahead.

MR. FRAWLEY: That idea I think it makes sense for us to say, look, this is actually a really important term for us. It's more important than term 55, we can get rid of that one. So it doesn't hurt their burden at all. It puts the onus on us to focus on what we think is important.

MR. GRATZ: And, your Honor, I just want to note that actually doesn't work in practice because we're not reviewing everything that hits on this search term, and then everything that hits on the next search term. We're getting a group of documents that hits on a group of search terms. But we're happy to work with the plaintiff to try to get them iteratively things that they want more, in hopes that they may not want the things that they want less.

THE COURT: Yeah. I mean, look, you all have managed to find so many ways to fight about so many things, that while I remain ever hopeful that this is the last time I'm going to hear about search terms, if I was a betting person, I wouldn't bet on myself.

Okay. That hopefully takes care of the search terms issues. I will review these other issues. You all have homework. Ms. Brook is standing up. I see another person on defendant's side standing up. Let's try to find a way to wrap up.

Go ahead, Ms. Brook.

MS. BROOK: This is in the direction of wrapping up. So two points. One, I'm sure this is already on the Court's agenda, but if we could set the next hearing before we leave today.

THE COURT: Absolutely.

MS. BROOK: Greatly appreciated.

And, two, as the Court noted, in our ability to find things to fight about, there are other issues that don't fall into the buckets of things that, as to both sides, that both sides have with each other's discovery to date, that don't fall into either something specifically raised with the Court today, or that we've specifically raised with the Court today. But the Court has indicated it will go back to the papers. Perhaps, in the Friday joint letters, we should just highlight the ECF entries associated with those issues that did not come up, or some other procedure that the Court prefers.

THE COURT: I had given you a larger homework assignment, which was not just --

MS. BROOK: Oh, the topic.

THE COURT: Yeah, the topics, and the chart by topics. Did I give you a due date on that?

MS. BROOK: I don't believe so, but I feel free to be corrected by 30 lawyers so.

THE COURT: Anyone?

MS. BROOK: My understanding was that was just how your Court wanted it done in advance of the next conference. So if it was more that your Court also wants one backwards --

THE COURT: Yes. Because I know there are things we haven't gotten to. I have a 30-page summary, prepared by my clerk, of the motions, and it's in tiny print too. So I know that there are issues that we haven't gotten to, but I'm asking

for your help to reframe those and point me to them. So that in addition to the bigger issues that you raised just now that we've been talking about and that you highlighted, they're in a chart or in buckets that I can more readily understand.

I think when we were talking about these issues today, I think it gives you some guidance about how I would like them framed, rather than by request number, but the broad issue of the dispute. And then what, you know, what it is or why this matters and then what the objections are.

MS. BROOK: May I suggest that we get that to the Court collaboratively by a week from today, the 6th of November?

THE COURT: One week is perfect. November 6th.

MS. BROOK: And then I'm going to sit down.

THE COURT: Okay.

MR. NATH: Your Honor, I have one question. We have joint letters due on two issues, I think if I counted it up correctly, on Friday. Does the Court want one joint letter or rather us file --

THE COURT: Two separate joint letters.

MR. NATH: Okay. Understood.

THE COURT: Yes. Introduce yourself, please.

MR. DAWSON: Your Honor, thank you. Your Honor, Andrew Dawson from Keker Van Nest & Peters. And my colleague Ms. Ybarra flagged on the docket we do have a motion, the

shorthand, which was damages. And I clearly heard your Honor's view that today is not the day for damages. I did want to make sure your Honor was aware, the damages issue highlighted in that motion are not just damages. They also relate to certain aspects of the fair use, the harm alleged in the complaint, product substitution. And even if it were just on damages, the defendants' concern at the moment is that the plaintiffs are, the New York Times in particular, are making productions about damages, but they're cabined in a way that are beneficial to the New York Times. And they have resisted the kind of discovery that we think is appropriate to our defenses.

So we're concerned about there being uneven playing field. And when experts start looking at materials, the only stuff in the record will be stuff the plaintiffs have decided is favorable to them, while they have resisted producing things that would be favorable to us. And I'm curious your Honor's view on how we should proceed about this, what timing your Honor would prefer, and the mechanism to bring these issues to the Court's attention.

THE COURT: This relates in some part to some of the issues raised in --

MR. DAWSON: So there's Dkt. 276.

THE COURT: Okay.

MR. DAWSON: And the docket number has three categories within it.

THE COURT: Whoa, whoa, whoa. Is that the one about like how New York Times use and views and statements about Generative AI?

MR. DAWSON: So there's some -- there are three categories. Two I think relate more to the pending motion from September. There's one that really is distinct, where the dispute at the moment is about analyses of trends. And the fundamental concern that we have is that it's no secret that The Times and Legacy Media are in a competitive marketplace. Lots of things influence the trends of subscription, of revenue, of website traffic.

The Times has been willing to produce materials that would -- within the category of AI's influence, but they're refusing to produce analyses that might be focused on entirely different trends, be they demographic trends, you know, competition from social media. I think that's the crux of the category for us.

If there are analyses out there suggesting that there was going to be a diminishment in subscriptions, totally independent of AI, totally independent of OpenAI, that's critical evidence for us. It's critical evidence on the product substitution issues that are relevant in fair use. And at some point, we do think we need to get access to that material so that the only information in the record is not the stuff most favorable to the plaintiffs.

THE COURT: Some of this relates to how analyses are done. So why don't you look at the analyses that are done first.

MR. DAWSON: And our concern is we don't know. We don't know what analysis have been done.

THE COURT: No, no. I'm saying what has been produced about the AI influence because, as I understand it, many of these analyses and data analysis, when drawing conclusions or inferences about trends, usually also rely on certain assumptions or other analyses that back out the effects of some of these other things that you're talking about. So I'm not sure that asking for separate analyses on these other issues is necessary if the analyses of the AI influence have already taken that into account.

MR. DAWSON: I think one concern we would have on that is timing. And this has come up, date range issues have come up more broadly. I won't get into those here. But I do think, I would assume that the trends within The Times and its leadership are complex, multivariable, lots of influences. The presence of AI as an alleged competitor, and to be clear, we think the allegations of substitution are kind of wildly speculative. The competition more broadly in the marketplace extends long before OpenAI came on the scene, long before Generative AI. And I must believe analyses from 2018, 2015 that go back before any of the new alleged influences came onto

the scene. And we think those are going to be highly valuable. And it could be that there are some more recent ones that take account of some of these multivariable. But we think if those more recent analyses end up not accounting for the same factors taken in by previous analyses, we need the full waterfront in order to assess and challenge the credibility of any allegations about the affect of OpenAI's technology.

THE COURT: Okay. What's the ECF numbers on that?

MR. DAWSON: So this is 276, and it's the first section within 276 is our motion. And the opposition is at number 293.

THE COURT: Okay. Put that in your chart that's coming in, in a week anyway. But I'll take a look at that also. Okay?

MR. DAWSON: Thank you, your Honor.

THE COURT: All right. So anything else you all want to raise before I have a few things that I was supposed to raise?

MR. TOPIC: Matt Topic for CIR. I have just some little housekeeping things.

THE COURT: Yeah.

MR. TOPIC: If you could enter the agreed PO and ESI order. Those are 103 and 104 on the CIR docket, that will then allow defendants to make their cross-production so we can start the process of getting caught up. I'm assuming it isn't going

to take them very long to do. That, you know, if you want to give them a date, I'd love that, but if you want to let it kind of ride, I'm okay with it.

THE COURT: What? To enter those protective orders?

MR. TOPIC: No, to make the cross-production to us. I don't want to wait weeks and weeks and months and months to get that.

THE COURT: So all you need is the protective orders entered and then they can cross produce.

MR. TOPIC: Correct.

THE COURT: That will be entered today. Within 24 hours, if not today.

MR. TOPIC: Thank you. I don't think we discussed it, but if we can get the interrogatory and RFA responses from the other cases, that will allow us to probably eliminate a whole lot of things we would otherwise be asking for.

THE COURT: Talk to your new friends.

MR. TOPIC: Thank you.

THE COURT: All right. We need to set a date for the next conference. I have a couple of other housekeeping things to go over with you, but I lost the post-it that they were written on.

But let's set the date. December. I think the first week of December. I had a trial set, so we have sometime.

THE DEPUTY CLERK: So any day that week. Tuesday, the

3rd.

THE COURT: Tuesday, December 3rd, 9:30 a.m., next conference.

Now, in preparation for the next conference, I will rely principally on the chart that you give me in a week. I will do my best to try to pare down that chart, but I'll plan to work primarily from that chart.

I have a few other housekeeping matters on my end as well. We no longer need hard copy courtesy copies. I know that's in my rules, but we have a room full of hard copy personal courtesy copies anymore. In a case like this, it's actually easier to rely on the electronic version.

The plaintiffs in The New York Times case were directed to file proposed exhibits to the amended complaint on ECF, and they've been filed ECF Nos. 171 through 202. Please also for that send to in electronic format to Judge Stein's chambers and my chambers. Confer with my deputy offline. I think the most expedient way may be via FTP. But we have to make sure we have it set up on our end to receive.

So no more hard copy courtesy copies. And let's see. I already asked you when you do file something, if it's an opposition or response, try to make at least one hyperlink to the other document. Just a minute.

I found my post-it. Okay. That's it on my end.

Yes, Ms. Brook, you had a question?

MS. BROOK: I may make one suggestion, your Honor.

THE COURT: Yeah.

MS. BROOK: So we have a due date from the Court already to submit the new chart in the format described by the Court by a week from today, November 6th. While the parties will do everything in their power to not add additional issues in between November 6th and December 3rd, I believe it was -- there may be one or two additional things. And so I was going to suggest that we provide the Court with an updated version of that chart in the format as described by the Court today. And so that associates and others are not doing that over the Thanksgiving week, I was going to suggest we make it due by the 22nd of November, which is the Friday before that Thanksgiving week, so a little bit more from a week out from that conference.

THE COURT: You read my mind. Updated chart, November 22nd. You either read my mind or you've had like conversations with other people about how I set due dates.

Not only is the updated chart due November 22nd, there is a moratorium on any filings about discovery. All I want between -- any new filings between November 23rd and the date of our conference are going to be administrative things, like *pro hac vice* motions. There is a moratorium on new filings. Okay. Please stop. And we'll use the updated chart from November 22nd as our agenda, as our guidepost for the December

conference.

At the end of the conference, maybe we can talk about whether there's new issues that are percolating or that may be ripe that arose between November 22nd and the date of the conference. But we're not going to argue them at that point. Then we're going to talk about how we're going to try to work them out, or whether we need briefing, letters or other topics.

Anything else?

Okay. My previous request about ordering the transcript stands. I can't remember, honestly, what I did last time. But thank you very much and we are adjourned.

(Adjourned)